JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2017JG004315

Assessing Microbial Community Patterns During Incipient Soil Formation From Basalt

Aditi Sengupta[1,2], James C. Stegen[2], Antonio A. Meira Neto[1,3], Yadi Wang[1,4], Julia W. Neilson[4], Triffon Tatarin[1], Edward Hunt[1], Katerina Dontsova[1,4], Jon Chorover[1,4], Peter A. Troch[1,2], and Raina M. Maier[4]

[1]Biosphere 2, University of Arizona, Tucson, AZ, USA, [2]Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA, [3]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA, [4]Department of Soil, Water and Environmental Science, University of Arizona, Tucson, AZ, USA

Abstract Microbial dynamics drive the biotic machinery of early soil evolution. However, integrated knowledge of microbial community establishment, functional associations, and community assembly processes in incipient soil is lacking. This study presents a novel approach of combining microbial phylogenetic profiling, functional predictions, and community assembly processes to analyze drivers of microbial community establishment in an emerging soil system. Rigorous submeter sampling of a basalt-soil lysimeter after 2 years of irrigation revealed that microbial community colonization patterns and associated soil parameters were depth dependent. Phylogenetic analysis of 16S rRNA gene sequences indicated the presence of diverse bacterial and archaeal phyla, with high relative abundance of Actinomyceles on the surface and a consistently high abundance of *Proteobacteria* (*Alpha*, *Beta*, *Gamma*, and *Delta*) at all depths. Despite depth-dependent variation in community diversity, predicted functional gene analysis suggested that microbial metabolisms did not differ with depth, thereby suggesting redundancy in functional potential throughout the system. Null modeling revealed that microbial community assembly patterns were predominantly governed by variable selection. The relative influence of variable selection decreased with depth, indicating unique and relatively harsh environmental conditions near the surface and more benign conditions with depth. Additionally, community composition near the center of the domain was influenced by high levels of dispersal, suggesting that spatial processes interact with deterministic selection imposed by the environment. These results suggest that for oligotrophic systems, there are major differences in the length scales of variation between vertical and horizontal dimensions with the vertical dimension dominating variation in physical, chemical, and biological features.

Key Points:
- A submeter sampling regime in an incipient homogenous basalt soil system revealed depth-dependent heterogeneous microbial presence
- Null modeling revealed strong influences of variable selection, resulting from localized abiotic environment selection pressures
- Length of scales of variation of hydrobiogeochemical signatures varied significantly along the vertical dimension but not horizontal

Supporting Information:
- Supporting Information S1
- Data Set S1
- Data Set S2
- Data Set S3
- Table S1
- Table S2

Correspondence to:
A. Sengupta,
aditi.sengupta@pnnl.gov

Citation:
Sengupta, A., Stegen, J. C., Meira Neto, A. A., Wang, Y., Neilson, J. W., Chorover, J., et al. (2019). Assessing microbial community patterns during incipient soil formation from basalt. *Journal of Geophysical Research: Biogeosciences*, *124*, 941–958. https://doi.org/10.1029/2017JG004315

Received 16 NOV 2017
Accepted 17 MAR 2019
Accepted article online 28 MAR 2019
Published online 12 APR 2019
Corrected 21 JUN 2019

This article was corrected on 21 JUN 2019. See the end of the full text for details.

Author Contributions:
Conceptualization: Aditi Sengupta
Data curation: Aditi Sengupta
Formal analysis: Aditi Sengupta, James C. Stegen
Funding acquisition: Jon Chorover, Peter A. Troch
(continued)

1. Introduction

The process of soil formation extends over a timescale of several thousand years. The five soil-forming factors (parent material, climate, topography, biota, and time) interact to define the state of a soil system and influence soil formation (Jenny, 1941; Weil & Brady, 2017). The biotic component is microbially driven in an incipient soil system lacking plants. In such incipient systems, pioneering microorganisms contribute a biological component to weathering, for example, by accelerating the oxidation of reduced metals in primary minerals. Additionally, microbes have acid-generating metabolisms and produce siderophores and other organic molecules that function as metal complexing agents to drive weathering reactions, thereby influencing the early buildup of labile nutrient pools that assist the growth of higher-order plant life (Bradley et al., 2014; Mapelli et al., 2012). These complex and coupled interactions provide a framework for soil genesis (Bockheim et al., 2014). However, these processes are difficult to observe in natural settings where initial soil conditions and time-variant forcings (e.g., water, energy and dust inputs, and parent material uniformity) are either unknown and/or cannot be controlled, making it challenging to accurately assess the contribution of the microbial community to soil formation.

Defining and measuring the drivers of microbial dynamics in an incipient system is key to understanding the coupling between soil formation factors during soil genesis. These drivers include in situ hydrological and geochemical processes. For example, structural properties of the subsurface will control infiltration and lateral redistribution of vadose zone water and groundwater and the resulting soil moisture spatial and temporal dynamics. These hydrologic processes determine the residence time of a parcel of water in contact with

Investigation: Aditi Sengupta, James C. Stegen, Antonio A. Meira Neto, Yadi Wang, Triffon Tatarin, Edward Hunt
Methodology: Aditi Sengupta, James C. Stegen, Antonio A. Meira Neto, Yadi Wang
Project administration: Jon Chorover, Peter A. Troch, Raina M. Maier
Resources: Raina M. Maier
Software: James C. Stegen
Supervision: Katerina Dontsova, Jon Chorover, Peter A. Troch, Raina M. Maier
Validation: Aditi Sengupta
Visualization: Aditi Sengupta, James C. Stegen
Writing - original draft: Aditi Sengupta
Writing – review & editing: Aditi Sengupta, James C. Stegen, Antonio A. Meira Neto, Yadi Wang, Julia W. Neilson, Jon Chorover, Peter A. Troch, Raina M. Maier

microbial cells and mineral surfaces undergoing dissolution and regulate the reactive transport of nutrients and microbes. Geochemical processes, including dissolution of primary minerals and precipitation of secondary phases, influence nutrient availability and may facilitate microbial colonization (Banfield et al., 1999). These drivers interact with the developing microbial community to influence the biotic-abiotic couplings that impact the soil genesis process.

Analysis of the assembly processes or forces that influence the specific combinations of taxa that comprise a microbial community is critical to understanding the dynamics of the developing microbial community. There is vast literature in ecology (e.g., Ackerly, 2004; Bazzaz, 1991; Bell, 2001; Cavender-Bares et al., 2004, 2009; Donoghue, 2008; Emerson & Gillespie, 2008; Enquist et al., 2002; Fine et al., 2006; Gillespie, 2004; Graham & Fine, 2008; He & Gaston, 2003; Hubbell, 2006; Jonathan Davies et al., 2007; Loreau, 2004; Mittelbach et al., 2007; Pennington et al., 2006; Ricklefs, 2004; Strauss et al., 2006; Swenson et al., 2007; Tilman, 2004; Vamosi et al., 2009; Webb et al., 2002; Weiher, 1999) focused on community assembly that provides insight into diversity profiles, assembly processes driven by microniche favorability in soil systems, and energy and metabolic requirements of microbial communities (Bradley et al., 2014; Caruso et al., 2011; Stegen, Fredrickson, et al., 2016). The ecological processes describe assembly of different microbial taxa that come together in a physical domain to constitute an intact community and range from deterministic to stochastic processes (Stegen et al., 2013). This work has shaped microbial community assembly studies (Burke et al., 2011; Dumbrell et al., 2010; Fan et al., 2012; Fuhrman, 2009; Green et al., 2004; Hanson et al., 2012; Horner-Devine et al., 2004; Nemergut et al., 2013; Pholchan et al., 2013; Sloan et al., 2007, 2006; Stegen, Fredrickson, et al., 2016; Woodcock et al., 2007), but few attempts have been made to examine the assembly patterns that arise in incipient soils (Brown & Jumpponen, 2014; Michel & Williams, 2011; Schmidt et al., 2014). This greatly limits understanding of the random stochastic and selective deterministic processes that influence the mechanisms of microbial community structures in rapidly evolving systems (Caruso et al., 2011), which in turn limits our capabilities to understand and predict the cumulative impact of microbial community dynamics on early soil development.

Here we assessed microbial community dynamics and assembly processes in an incipient basaltic soil system. It is noted that basalt in seafloors, soda lakes, and terrestrial environments has been known to host endemic endolithic microbial communities (Bengtson et al., 2014; Cockell et al., 2009; Orcutt et al., 2015; Singer et al., 2015; Wani et al., 2006). However, there is no previous research known to the authors that provides a view of incipient microbial community structure and assembly in basaltic porous media subjected to through-fluxes of fresh meteoric water. Here microbial community dynamics and assembly processes were assessed in a mesocosm-containing fresh granular basalt, termed *miniLEO*, a lysimeter with dimensions ($L \times D \times W$) of 2 m × 1 m × 0.5 m, and a slope angle of 10° housed at University of Arizona's Biosphere 2 facility. The mesocosm (Sengupta et al., 2016) has low environmental and trophic complexity and is designed to replicate the three massive zero-order basin hillslopes of the Landscape Evolution Observatory (LEO) at Biosphere 2 (Pangle et al., 2015). The bottom boundary of the lysimeter is impermeable (no flow boundary) and all water effluxes through the seepage face (Kim et al., 2016; Sengupta et al., 2016). Following a 2-year period of controlled experimentation (including initially homogenous parent material, regimented irrigation, two-dimensional flow system with no significant flow along the width, and a climate-controlled environment), miniLEO was excavated in a spatially intensive manner. Phylogenetic 16S rRNA gene amplicon sequencing was performed to determine patterns of bacterial and archaeal relative abundance and diversity and assess whether significant spatial heterogeneity of microbial taxa could be observed in the initially homogeneous basalt system. System properties (bulk density, electrical conductivity, moisture content, pH, and carbon and nitrogen content) were quantified to identify explanatory variables driving observed variances in microbial community diversity and composition. PICRUSt (Langille et al., 2013) was then used to generate preliminary hypotheses for microbial functional potential and its spatial heterogeneity in the incipient system. The classical expectation that shifts in community composition are governed deterministically by variation in environmental factors does not account for spatial processes, stochasticity (Stegen et al., 2012), or ecologically neutral processes (Rosindell et al., 2012) that may influence community composition. Therefore, ecological null models (Stegen et al., 2012, 2015) were used to estimate the relative influence of deterministic ecological selection and to further parse the influences of variable and homogenous selection and dispersal limitation and homogenizing dispersal.

2. Materials and Methods

2.1. Experimental Design and Sampling Scheme

Samples were collected from basaltic mesocosm (hereafter referred to as miniLEO) with dimensions of 2 m length, 0.5 m width, 1 m height, and slope of 10° (Figure 1a) housed within the Landscape Evolution Observatory (LEO) in Biosphere 2 (B2) at the University of Arizona. Basaltic tephra, mined from the Merriam Crater near Flagstaff, AZ, crushed to a loamy sand texture (hereafter referred to as soil material) served as the homogenous parent material. Detailed description of particle-size distribution, mineral composition of the parent material, and instrumentation is provided in Pangle et al. (2015, 2017). The system was subjected to 15.2 m (14,440 L) of irrigation with reverse-osmosis filtered water over a period of 644 days during years 2014–2015 (Figure 1c). Irrigation was delivered as controlled pulses throughout the experiment (Supporting Information Figure S1), separated by dry periods. During the entire duration of the experiment miniLEO was exposed to the internal environment of the enclosed LEO space although we cannot eliminate the possibility of air inputs.

An intensive excavation and sampling protocol was executed over a period of 2 weeks in September 2015 (Sengupta et al., 2016). Briefly, the mesocosm dimensions were divided into a Euclidean space of X, Y, and Z coordinates, where X denotes width, Y denotes length, and Z denotes depth (Figure 1b). Next, X was divided into four intervals of 10 cm each ($X = 0$, 1, 2, and 3), Y was divided into 10 intervals of 20 cm each ($Y = 0$ to 9), and Z was divided into 10 intervals of 10 cm each ($Z = 0$ to 9).

Each unique combination of XYZ was identified as a $10 \times 20 \times 10\text{-cm}^3$ voxel, which resulted in 400 unique samples that could be collected from miniLEO. During the excavation process, the last regions along length and depth (Y and $Z = 9$, 76 samples) were discarded to eliminate sample contamination from preceding layers. Additionally, a boundary of 5 cm along the walls and floor and an additional $0.1 \times 0.5 \times 1\text{-m}^3$ gravel layer at the toe slope were discarded. This resulted in 324 samples ($X = 0$ to 4, $Y = 0$ to 8, $Z = 0$ to 8) collected. A soil corer of 15 cm × 0.5 cm was used to collect 10-cm cores for microbiological samples from each voxel as it was collected. The corer was flame sterilized prior to each collection. Samples were deposited in sterile pouches, and they were stored at −80 °C until DNA extraction. Following microbiological sampling, samples from each voxel were also collected for physicochemical measurements analysis including pH (U.S. EPA method 150.2), electrical conductivity (EC; U.S. EPA method 120.1), total carbon (TC), inorganic carbon (IC), organic carbon (OC), and total nitrogen (TN; U.S. EPA method 415.3), and bulk density (BD; Supporting Information Data Set S1). The distribution of spatially distinct moisture regimes was captured by modeling data from 13 spatially distributed water content sensors for dry and wet periods of 2014 and 2015 (herein referred as WC_2014_Dry, WC_2014_Wet, WC_2015_Dry, and WC_2015_Wet). Figure 1b illustrates the wet (irrigation) and dry (no irrigation) periods for each year, with irrigation spells signifying wet periods, interspersed with dry periods that saw no irrigation. Since samples were collected in 2015, average water content in 2015 (WC_Av_2015) was also included in the analysis. Each sensor time series was averaged for the chosen period and then spatially interpolated to provide voxel-based results.

2.2. DNA Extraction

DNA was extracted from 2 g of sample (voxel cores, parent material) using the Fast DNA SPIN for Soil Kit™ (MP Biomedicals, Solon OH, USA) with modifications to the manufacturer's protocol. A total of 162 samples were chosen, representing voxels from $X = 2$ and 3, and all Y and Z locations. The samples from $X = 0$ and 1 voxels were not used since most sensors were positioned in these regions, and they occasionally did not have intact sample integrity for parallel geochemical and hydrological analyses. Additionally, DNA was extracted from 10-ml irrigation water and from extraction blanks. All consumable and kit supplies, except those that have biomolecules, were UV-sterilized for 30 min. Soil samples were thawed on ice prior to extraction. Since the soil material was highly oligotrophic, a combined extraction step was performed during the initial stages of extraction, with 0.5-g soil added to each of four extraction tubes per sample. Supernatants from each of the tubes were pooled during the Binding Matrix-DNA binding step by sequentially adding DNA supernatant from the individual tubes and rinsing the composite supernatant 2x with Binding Matrix supernatant. Additionally, the spin filters containing the binding matrix were air-dried under laminar flow hood for 10 min prior to DNA elution, followed by elution with UV-sterilized water preheated to 60 °C. The extracted DNA was quantified using the Qubit® dsDNA High Sensitivity Assay

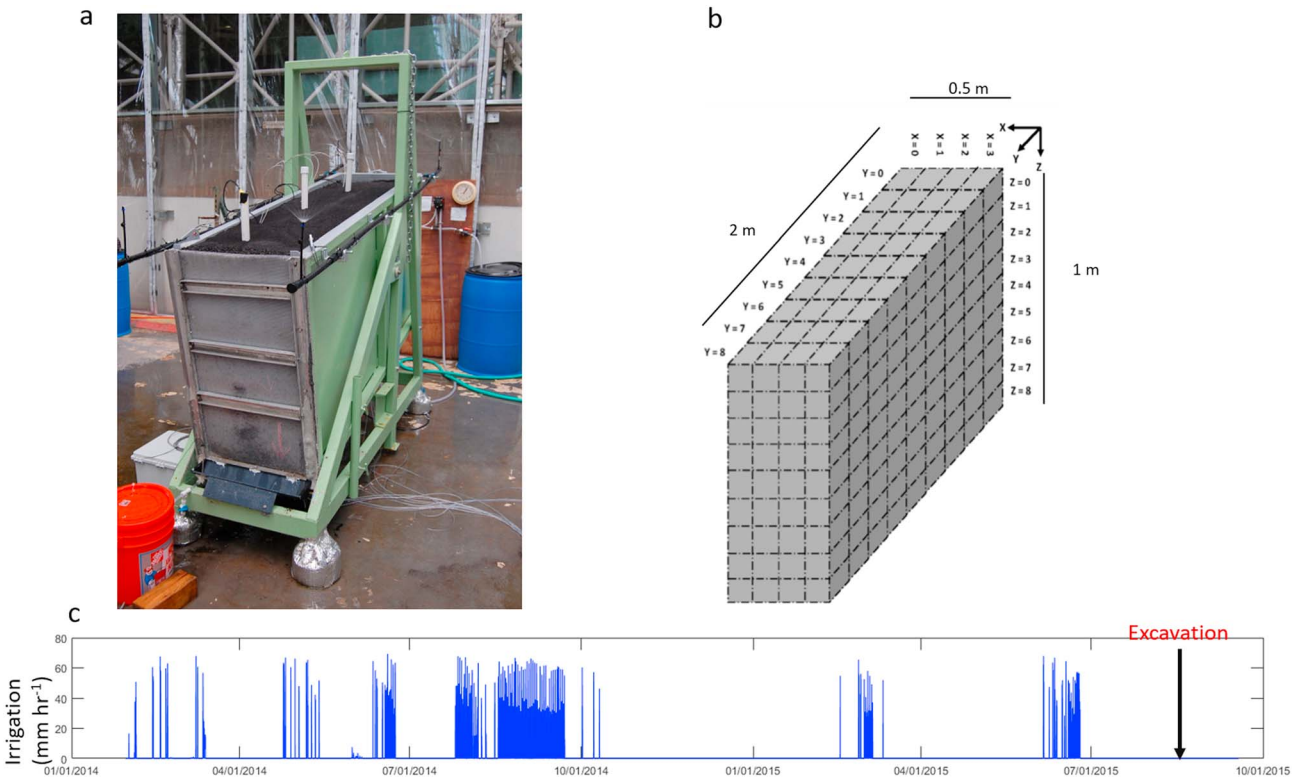

Figure 1. (a) miniLEO lysimter filled with crushed basaltic tephra. (b). *X*, *Y*, *Z* dimensions of voxels. (c) Irrigation rates applied (mm/hr).

Kit (Life Technologies, NY, USA) with those falling below the minimum detectable threshold of 0.01 ng/ul identified as below detection limit.

2.3. Illumina Sequencing

For each of the 162 samples, parent material, irrigation water, and extraction blanks, paired-end sequencing (2 × 150 bp) was performed on the bacterial and archaeal 16S rRNA gene V4 (515F-806R primers) hypervariable region using the Illumina MiSeq platform (Illumina, CA, USA; Caporaso et al., 2012). All of the sequencing procedures, including the construction of Illumina sequencing libraries, were performed using the protocol previously published (Caporaso et al., 2012) with modifications (Laubitz et al., 2016). Illumina MiSeq v2 (300 BP) chemistry was used for sequencing and was performed on the Illumina MiSeq (SN M02149, with the MiSeq Control Software v 2.5.0.5) at the University of Arizona Genetics Core (Tucson, AZ, USA) following their standard protocols. The Genetics Core provided standard Illumina quality control, base calling, demultiplexing, adaptor removal, and conversion to FastQ format. Raw sequence data were submitted to NCBI's Sequence Read Archive SRP116044 (Accession PRJNA392820), and sequence information per sample is provided in Supporting Information Data Set S2.

2.4. Bioinformatics Analyses

Paired-end sequence merging, barcode removal, quality filtering, singleton-sequence removal, chimera checking and removal, and open-reference Operational Taxonomic Unit (OTU) picking were conducted using default parameters unless otherwise specified in Qiime v 1.9.1 (Caporaso et al., 2010). A minimum of 20 base overlap was specified for joining the paired reads to give an average sequence length of 253 base pairs. A summary of the sequences, postmerging and quality filtering, was performed using mothur (v 1.25; Schloss et al., 2009), OTU picking was done using UCLUST (Edgar, 2010), and sequence alignment was performed with PyNAST (Caporaso et al., 2010). Clustering was done with Greengenes database at 97% sequence similarity (DeSantis et al., 2006), chimera were removed with Chimera Slayer (Haas et al., 2011), taxonomy was assigned with RDP Classifier (Wang et al., 2007), tree building was completed with FastTree (Price et al., 2010), and comparative diversity calculations were done with UniFrac (C. Lozupone

& Knight, 2005). OTUs that were observed only once after chimera filtering were removed. The OTUs from 162 miniLEO samples were grouped according to their width profile ($X = 2$ and $X = 3$) and compared for community diversity (dis)similarities between samples located along the X dimension of miniLEO. This approach was taken to inform downstream analyses of a subset of samples. Processed sequence information is provided in Supporting Information Data Set S2.

Multiple rarefactions were performed at sequence depth of 5,000–25,000 and collated for alpha diversity estimates including Shannon (H`), richness (observed OTUs), and phylogenetic diversity. To determine diversity between samples (beta diversity), weighted Unifrac distance (C. Lozupone & Knight, 2005) was calculated to account for abundance and phylogeny. All data files generated from QIIME workflow were imported into *R* environment program (R Core Team, 2014) for alpha and beta diversity estimation and visualization using *Phyloseq* (McMurdie et al., 2013), statistical analyses using *vegan* (Oksanen, 2015), and heatmap generation using *gplot* (Warenes et al., 2016).

2.5. Null Modeling

We applied a previously developed null modeling framework (Stegen et al., 2013, 2012) to estimate the relative influences of community assembly processes. The framework estimates the degree to which community composition in one sample differs from the other samples as a result of different ecological processes (homogeneous selection, variable selection, homogenizing dispersal, and dispersal limitation; Dini-Andreote et al., 2015; Stegen et al., 2015). Homogeneous selection occurs when the environment constrains communities to be very similar, and variable selection occurs when the environment drives community composition apart (Dini-Andreote et al., 2015). Variable selection occurs when selective pressure in environmental conditions cause turnover of species. Dispersal limitation occurs when there is very little exchange of organisms between a given pair of communities and when selection-based processes do not strongly drive the communities to be similar or different. Under these conditions, stochastic birth/death events lead to ecological drift and divergence in community composition. Homogenizing dispersal occurs under similar conditions except that there is a high rate of organismal exchange between a given pair of communities.

The methods used to infer the relative contributions of these assembly processes have been detailed previously (see citations above); here we provide a brief summary. The null modeling approach starts with community composition and inferred phylogenetic relationships among OTUs. Using those data, we calculated the between-community mean-nearest-taxon-distance (βMNTD) metric (Stegen et al., 2012) for each pairwise community-to-community comparison. For each pairwise comparison, a null distribution of βMNTD values was then generated. This was done by randomizing phylogenetic relationships among OTUs and recalculating βMNTD and repeating this procedure 999 times to generate a null distribution of βMNTD values. For each pairwise comparison, homogeneous selection or variable selection was inferred as the ecological basis of community dissimilarity if the observed βMNTD value was significantly less or greater than the null distribution, respectively. The β-Nearest Taxon Index (βNTI) was used to evaluate significance. βNTI expresses the difference between observed βMNTD and the mean of the null distribution in units of standard deviations with βNTI values <-2 or $> +2$ indicating significance.

If observed βMNTD does not significantly deviate from the null expectation, then the observed compositional difference is not due to selection and may be due to either homogenizing dispersal or dispersal limitation. In these cases we used a version of the Raup-Crick metric that accounts for OTU relative abundances, known as RC_{bray} (Stegen et al., 2013). Like βNTI, RC_{bray} compares observed dissimilarity to a null distribution but uses the Bray-Curtis dissimilarity metric instead of βMNTD. For each pairwise comparison the null Bray-Curtis distribution was generated using 999 null model runs that simulated stochastic community assembly (see Stegen et al., 2013, for details). A value of $RC_{bray} > +0.95$ indicates greater dissimilarity than expected and when paired with a βNTI value that is nonsignificant (i.e., $|\beta NTI| < 2$), a dominant influence of dispersal limitation is inferred. Similarly, $RC_{bray} < -0.95$ indicates communities are more similar than expected and when paired with $|\beta NTI| < 2$, a dominant influence of homogenizing dispersal is inferred. If, for a given pairwise comparison, neither null model is significant (i.e., $|\beta NTI| < 2$ and $|RC_{bray}| < 0.95$), we infer that the observed dissimilarity was not the result of any one process, and this situation is referred to a being *undominated* (Stegen et al., 2015). R code for running the null models can be found here: https://github.com/stegen/Stegen_etal_ISME_2013.

2.6. Statistical Analyses

2.6.1. Comparison of Samples From $X = 2$ and $X = 3$

The OTU frequencies in sample groups ($X = 2$ and 3) were compared using a nonparametric T test with $n = 1{,}000$ Monte Carlo simulations to determine statistically significant differences between the OTU abundances in the groups. Additionally, samples were plotted as ordinations of weighted Unifrac distance to capture variation in the (dis)similarity of the communities.

2.6.2. Environmental Variation

ANOVA was used to determine significant environmental variables with respect to categorical variables (depth and length). This was followed by analysis of covariance (ANCOVA) of environmental variables by fitting linear models to understand which depth and length groups best explained variation in the environmental variables. ANCOVA, which is a combination of ANOVA and regression, was performed 999 times on each variable to explain covariances observed for the variables as a function of depth and length. For all tests, P0.05 was considered a statistically significant difference.

2.6.3. Using Environmental Data to Explain Diversity Data

Samples were grouped into categorical variables as depth dependent (Surface, Depth_10, Depth_20, Depth_30, Depth_40, Depth_50, Depth_60, Depth_70, and Depth_80) and length dependent (TopSlope1, TopeSlope2, TopSlope3, MidSlope1, MidSlope2, MidSlope3, ToeSlope1, ToeSlope2, and ToeSlope3). Alpha diversity statistics were evaluated using ANOVA to determine depth- and length-dependent significance. Weighted Unifrac measure was calculated based on log-transformed relative abundances of OTUs and differences in community structure (β-diversity) displayed with principal coordinate analysis (PCoA) ordinations. ANalysis of SIMilarities (ANOSIM) and r statistics were used to test whether two or more groups of samples were significantly different based on a categorical variable. Adonis function, which partitions a distance matrix among sources of variation and performs permutation multivariate analysis, was used to test the strength and significance of measured (BD, pH, EC, TC, TIC, TOC, and TN) and modeled variables (water content) on Bray-Curtis community distance matrix. Significance tests were conducted using F tests based on sequential sums of squares from permutations of the raw data. The measured and modeled metadata are provided in Supporting Information Data Set S1. Spearman Rank Correlations were conducted to determine significant correlation of dominant phyla in miniLEO to measured and modeled environmental variables. A robust correlation was considered if the Spearman's correlation coefficient was >0.4 and AdjPValue0.05. Additionally, regression analyses of βNTI values with changes in environmental variables were performed.

3. Results

3.1. Low-Template DNA Samples and Selection of Community Comparisons

The DNA concentration of the samples ranged from those below detection limit of (<10 pg/g) to a maximum of 20-ng/g dry weight of soil. Based on OTU abundances in $X = 2$ and 3, and the lack of evidence of water flow in the X direction, a single width block ($X = 2$, 80 samples) was analyzed downstream. Additional sample selection information and community characterization of parent material and irrigation water is provided in Supporting Information S1. Briefly, five and four phyla were greater than 1% in parent material and irrigation water, respectively. The parent material was dominated by *Bacteroidetes* (60%) followed by *Firmicutes* (15%), while irrigation water had high relative abundance of *Bacteroidetes* (80%) with *Verrucomicrobia* being the second dominant phyla (18%). The relative abundance of *Proteobacteria* and *Actinobacteria* was <5% in both parent material and irrigation water.

3.2. Submeter-Scale Variation of Environmental Variables

A 200-cm^3 voxel was the sampling unit used to examine heterogeneity of environmental variation in the miniLEO, focusing on the two-dimensional slice through the mesocosm at $X = 2$. The measured environmental variables on these 80 voxels and the modeled water content variables calculated for 62 voxels (Supporting Information Data Set S1) were analyzed for depth-dependent and length-dependent associations. Significant variation was observed in BD, EC, TC, IC, TN, and WC with respect to depth, while length significantly affected the pH and WC during dry periods (Table S1a). Analysis of covariances showed heterogeneous relationships between the significant variables and their location based on whether they were grouped along the depth profile or the length profile. Table S1b lists the significance of each depth and

length group on measured and modeled variables. The pH of the miniLEO samples was significantly lower (median = 9.29) than the parent material (10.1); however, the pH did not significantly change with depth. Instead, significant differences were observed with respect to length. For each depth except 0–10-cm samples, the pH dropped from the top slope (average 9.3) to the midslope region (average 9.2) and then increased gradually at the toe-slope region (average 9.4). Bulk density was significantly higher in the last layer (average 1.5 g/cm) as compared to the other layers (average range 1.3–1.4 g/cm), while EC drastically dropped from 400 μS/cm (±105 μS/cm) in the surface to less than 100 μS/cm (±13 μS/cm) on average in the second layer and below. Trends similar to EC were observed for TC, IC, and TN, where surface samples were significantly higher than the other layers. Despite observing a significant increase in organic carbon in the miniLEO samples (average 0.14 ± 0.09 mg/g) from the original parent material (9.33×10^{-5} mg/g), the organic carbon profile did not significantly differ between the layers, rather the total carbon profile was influenced by the inorganic carbon concentrations.

The total carbon (0.28 ± 0.09 mg/g) and inorganic carbon (0.18 ± 0.05 mg/g) concentration in the surficial layer was significantly greater than the average total carbon (0.12 ± 0.07 mg/g) and inorganic carbon concentration (0.05 ± 0.02 mg/g) in layers 10–80 cm. Total nitrogen concentrations were significantly higher in the miniLEO samples (0.007 ± 0.004 mg/g) than the parent material (4.33×10^{-6} mg/g), with the surficial samples averaging significantly higher concentrations (0.018 ± 0.004 mg/g) than the 10–80-cm samples (0.006 ± 0.002 mg/g).

Water content variation was also related to depth profiles during the dry periods with an average increase in the content for samples located 10 cm and below. The 70–80-cm samples averaged twice the amount of water content as the surficial layer. Wet periods resulted in significant gradual increase in average content with depth (22.82 ± 1.90 in 0–10-cm to 36.41 ± 2.19 in 70–80-cm samples). Overall, the total water content of each voxel was 1.5–2x times higher during the wet period than dry. Water content along the length profiles showed significant increase in the toe slope as compared to top and middle in the dry periods and an overall increase from top to toe slope during dry and wet periods. Since the mesocosm was dry for a greater part of the experiment in 2015, WC_Av_2015 followed patterns of 2015 dry periods for length- and depth-dependent variations.

3.3. Community Composition

A total of 5,199,080 bacterial and archaeal sequence reads were obtained from 80 samples, resulting in 8,037 OTUs. OTUs that had more than 0.01% total sequences per sample were retained for further analysis. The top 10 phyla included *Proteobacteria* (29.8%), *Euryarcheota* (16.7%), *Actinobacteria* (14.1%), *Planctomycetes* (7.7%), *Verrucomicrobia* (5.1%), *Chloroflexi* (4.8%), *Nitrospirae* (3.6%), *Bacteroidetes* (3.4%), OD1 (3.3%), and *Acidobacteria* (3%). The remaining 30 phyla represented 8% of the communities. Among proteobacterial OTUs, *Betaproteobacteria* was the most abundant class (9.5%) followed by *Alphaproteobacteria* (9.3%), *Gammaproteobacteria* (7.4%) and *Deltaproteobacteria* (3.2%). Relative abundances were summarized along the depth and length profiles of the lysimeter. The greatest difference in terms of depth was between the surface 10-cm layer and the samples at 80-cm depth (Figure 2). Actinobacterial OTUs were predominantly greater at the 0–10-cm samples, while proteobacterial OTUs were relatively more abundant at the 80-cm depth. The Actinomycetales order, known to form filaments and hyphae, was abundantly found in miniLEO, with the surficial samples accounting for 13% as compared to the deeper layers averaging 2.7%. When compared to the surficial samples, depth-dependent increases were observed for *Verrucomicrobia*, *Euryarchaeota*, and *Bacteroidetes*, while *Planctomycetes*, *Nitrospirae*, and *Chloroflexi* showed a more consistent pattern. Relative abundances of OTUs along the length profile were not significantly different (Supporting Figure S4). The parent material and irrigation water had high relative abundance of *Bacteroidetes*, while *Proteobacteria* and *Actinobacteria* were less abundant as compared to the miniLEO samples. The parent material also had higher relative abundance of *Firmicutes* as compared to irrigation water whereas *Verrucomicrobia* was relatively abundant in irrigation water. Archaeal OTUs were not observed in the parent material and irrigation water.

3.4. Microbial Community Structure and Diversity

Alpha diversity statistics including observed richness, Shannon diversity, and Faith's Phylogenetic Diversity are provided in Table S2. One-way ANOVA showed significant differences in observed richness of samples

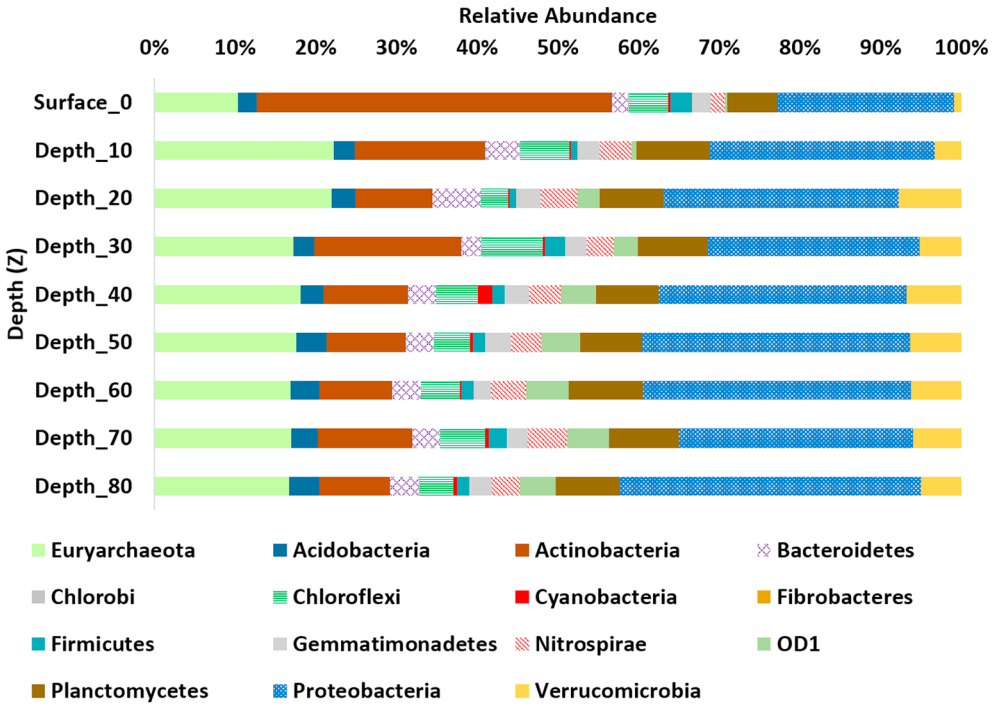

Figure 2. Relative abundance depth profile of top 15 phyla in miniLEO based on 16S rRNA gene-amplicon sequence.

grouped according to their depth (p_{depth} = 0.004) but not length profiles (p_{length} = 0.55), with richness peaking at 10–20-cm depth, gradually decreasing with depth, and increasing again in the 80–90-cm samples. Diversity metrics revealed only Faith's Phylogenetic Diversity (Faith's PD) to be significantly different between samples grouped according to their depth ($p_{Faith'sPD\_Depth}$ = 0.008, $p_{Faith'sPD\_Length}$ = 0.37; $p_{Shannon\_Depth}$ = 0.07, $p_{Shannon\_Length}$ = 0.16), with average Faith's PD increasing to 127.1 in the 10–20-cm samples, decreasing gradually to 88.2 in the 50–60-cm samples and increasing again to 127 in the 80–90 samples. A weighted UniFrac metric comparison using PCoA ordination showed a gradient in community composition as a function of depth. The surface-10 cm formed a distinct community followed by an incremental progression from 10–20 to 20–30 cm then all samples from 30 to 90 cm (Figure 3). Clustering patterns were not observed when samples were analyzed according to their

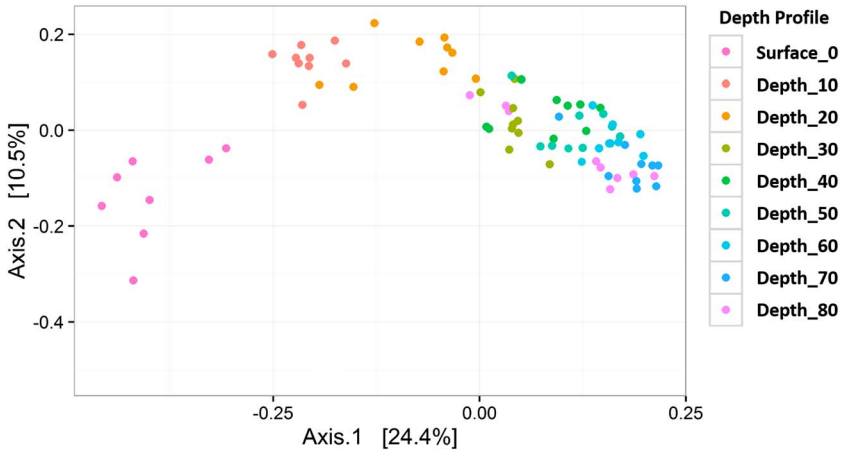

Figure 3. Principal coordinate ordination of weighted Unifrac distance of samples in miniLEO, grouped according to depth.

length profiles (Supporting Information Figure S5). A Similarity Percentage Analysis using agglomerative hierarchical clustering of Bray-Curtis distance showed samples broadly clustering into two groups, with surface samples distinctly separate from the rest (Supporting Information Figure S6).

Adonis tests on Bray-Curtis distance matrix with the measured and modeled variables explained the variability in β-diversity and constrained variance in the microbial population. The significant explanatory variables included BD, EC, TC, IC, TN, WC_2014_Dry, WC_2015_Dry, WC_2014_Wet, WC_2015_Wet, and WC_Av_2015 ($p = 0.001$). Next, phylogenetic composition of the samples was evaluated to identify correlations between the top 10 abundant phyla-level taxa and the significant explanatory variables (Supporting Information Figure S7). *Proteobacteria* was negatively correlated with TN ($r_s = 0.41$, $q = 0.004$) and positively correlated with WC_Av_2015 ($r_s = 0.42$, $q = 0.004$). *Euryarcheota* was negatively correlated with EC ($r_s = 0.41$, $q = 0.004$), whereas *Actinobacteria* was positively correlated with TN ($r_s = 0.48$, $q = 0.00004$) and IC ($r_s = 0.41$, $q = 0.0005$), but strongly negatively correlated with WC_Av_2015 ($r_s = 0.58$, $q = 0.000003$). *Verrucomicrobia* was negatively correlated to TN ($r_s = 0.42$, $q = 0.002$) and IC ($r_s = 0.42$, $q = 0.002$). *Acidobacteria* was negatively correlated with TN ($r_s = 0.49$, $q = 0.00002$) and IC ($r_s = 0.43$, $q = 0.003$), but positively correlated with WC_Av_2015 ($r_s = 0.44$, $q = 0.0005$), whereas *Bacteroidetes* was negatively correlated with TC ($r_s = 0.41$, $q = 0.002$). OD1 were strongly negatively correlated to TN ($r_s = 0.59$, $q = 0.00000004$) and IC ($r_s = 0.54$, $q = 0.0000007$) and positively correlated to WC_Av_2015 ($r_s = 0.61$, $q = 0.0000005$). *Planctomycetes*, *Chloroflexi*, and *Nitrospirae* did not show any significant correlation with the variables. Therefore, it appears that microbial response to environmental variables was heterogeneous and that phylum-environment relationships in miniLEO were not homogeneous.

The depth-dependent dissimilarity (Figure 3) and heterogeneous microbial community composition motivated us to evaluate microbial functional potential of miniLEO and generate preliminary functional hypotheses for the oligotrophic basalt system. Due to the oligotrophic nature of the system, we chose to focus primarily on carbon and nitrogen cycling functions. Briefly, using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) pipeline (Langille et al., 2013), a closed-reference OTU composition was normalized by the 16S rRNA gene copy number to predict metagenome functional content inferred using the Kyoto Encyclopedia of Genes and Genomes (Kanehisa et al., 2017). From a predicted 6,900 Kyoto Encyclopedia of Genes and Genomes descriptions, the average log-transformed abundance counts of select few carbon and nitrogen metabolism functions were compared across surficial (0–10 cm), Depth_10 (10–20 cm), Depth_20 (20–30 cm), and Depth_30_80 (30–80 cm) samples.

Enzyme categories associated with carbon fixation (Ribulose bisphosphate carboxylase catalytic and N-terminal domain, *cbbM* gene) were predicted to be uniformly high in the mesocosm. The commonly associated orders of microbial taxa (Rhodobacterales, 0.4%; Acidimicrobiales, 5.8%; and Nitrospirales, 3.7%) affiliated with these gene families were uniformly distributed in our 16S rRNA data set. The key enzyme for reverse TCA cycle, citrate lyase, could not be predicted though counts of other TCA cycle enzymes (methylisocitrate lyase and isocitrate lyase) were predicted (Supporting Information Figure S8). These genes were of interest due to their presence in seafloor basalt (Orcutt et al., 2015). An alphaproteobacterial methanotroph, *Methylocystaceae* (0.1%), was also observed across the depth profile.

The nitrogen metabolism pathways presented an interesting observation. While proteobacterial order Rhizobiales (4.4%) known to fix nitrogen and Nitrospirae (3.7%) phyla known to play a role in nitrification were classified in the OTU table, PICRUSt did not predict gene families related to nitrogen fixation (nitrogenase, *nif*H), nitrification (ammonia monooxygenase subunit C, *amoAC*), hydroxylamine oxidase (*hao*), nitric oxide reductase (*norA*), and anaerobic ammonia oxidation (hydrazine hydrolase, *hzo* and hydrazine dehydrogenase, *hdh*). Dissimilatory nitrite reductase (denitrification; *nir*K, *nir*B) and nitrate reductase (dissimilatory nitrate reduction, denitrification; *nar*G, H, J) predictions were uniformly abundant in all four depth groups (Supporting Information Figure S8) with these functions found in *Deltaproteobacteria* (Bu et al., 2017; Nelson et al., 2016; Slobodkina et al., 2017; Thorup et al., 2017), which were uniformly distributed in miniLEO (ranging from 2.1% to 4.0% across the depth profile). Thus, the depth-dependent variability in microbial community diversity and community composition did not impact predicted carbon and nitrogen cycling traits.

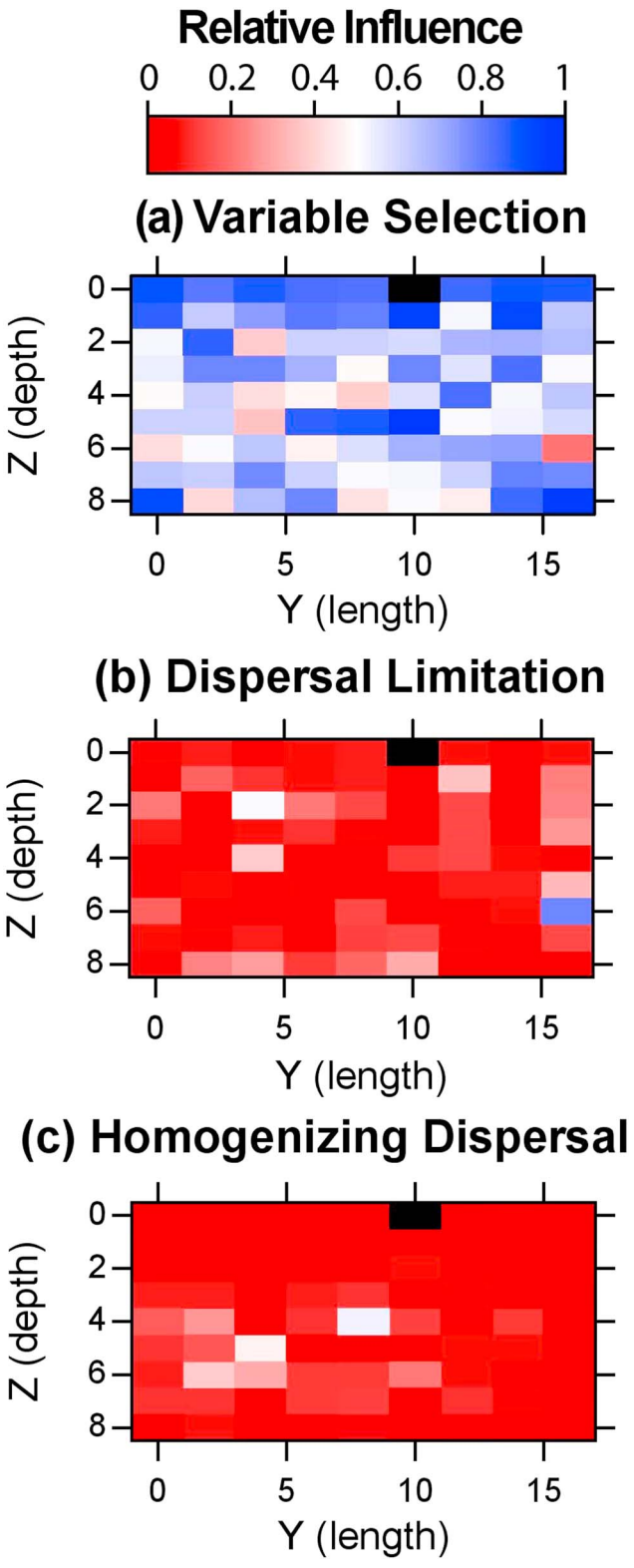

Figure 4. Spatial structure of community assembly processes, showing the relative contributions of (a) variable selection, (b) dispersal limitation, and (c) homogenizing dispersal. *Y* axis is *Z* (depth) increments of 10 cm each starting at 0 for surface-10-cm below samples. *X* axis is *Y* (length) increments of 20 cm each starting at 0 for Toe-Slope1 samples.

3.5. Community Assembly

The system as a whole was dominated by variable selection, which was the primary driver of between-community shifts in composition for ~66% of the pairwise comparisons (Supporting Information Figure S9). The spatial patterns in variable selection, dispersal limitation, and homogeneous dispersal across miniLEO, respectively, are presented in Figures 4a–4c. To generate these panels, we estimated the degree to which each assembly process led to differences in composition between the community found within a given voxel and all other communities. This was done by examining the null model outputs for all pairwise comparisons involving the community found within a given voxel. In Figure 4a, colors indicate the relative influence of variable selection on the community in each voxel (i.e., the fraction of pairwise comparisons that resulted in βNTI > +2). Dark blue voxels indicate that the associated community is differentiated from most other communities in miniLEO because it has a unique environment that deterministically causes it to be distinct (i.e., variable selection is strong). Dark red indicates that the community in that voxel is not differentiated from other communities by variable selection. The midpoint is shown in white, which indicates that the community is differentiated by variable selection from approximately half of the communities found throughout miniLEO.

Figure 4a reveals that while variable selection is the dominant assembly process in miniLEO, its influence changes significantly across voxels. This spatial variation was characterized by a relatively weak, yet significant, vertical gradient whereby the influence of variable selection was maximized at the top of miniLEO (Supporting Information Figure S10a).

As noted above, variable selection governed 66% of community pairwise dissimilarities, and of the remaining 34%, dispersal limitation and homogenizing dispersal were responsible for 11% and 7%, respectively. The effects of dispersal limitation were spatially variable but showed no clear patterns (Figure 4b). Homogenizing dispersal, however, appeared to follow a unimodal pattern across the vertical dimension (Supporting Information Figure S10b). Locations near the extreme ends of the horizontal axis also appeared to have very low influences of homogenizing dispersal (Figure 4c). Regression relationships between βNTI values and environmental variables (Supporting Information Figure S11) showed that most likely multiple variables were affecting the spatial turnover in community composition, with similar results for TN, IC, and EC.

4. Discussion

The controlled experimentation (regimented precipitation, semiclosed surroundings, and absence of plants) of a homogeneous soil system supported revealed that spatial heterogeneity of microbial communities in an oligotrophic basalt soil system was a function of depth. Variable selection was the primary driver of the community assembly, which was found to be especially strong in the surficial layers. Notably, microbial communities in the surficial layers were exposed to higher salt, lower moisture, greater amplitude fluctuations in temperature, greater interaction with the atmosphere, and higher nitrogen content from irrigation water. The elevated influence of variable selection in the surficial layers indicates that these environmental conditions imposed strong selective pressures that resulted in ecologically deterministic assembly of

microbial communities. These heterogeneous community patterns suggest that ecological selection pressures were strong and spatially variable and that localized selection pressures were particularly unique near the surface of miniLEO. Depth-dependent soil microbial characteristics have been revealed in well-developed soils (Gittel et al., 2014; Hansel et al., 2008; Hartmann et al., 2009; Seuradge et al., 2017; Will et al., 2010). We suggest that the stage is set early in the history of soil development where deterministic selection pressure drives soils to exhibit depth-dependent microbial community profiles and are compounded overtime as the system experiences additional selection pressure from plant growth and vegetation input.

Our results suggest that at early stages of soil development, localized environmental pressure selects for and shapes the early microbial patterns. An exciting addition to this concept would be understanding lineage-specific age-related shifts in microbial communities. A recent study (Turner et al., 2017) revealed that long-term soil development differentially affected bacterial and archaeal communities, the latter better adapted to subsoil environment. Therefore, conceptually, one can hypothesize that bacterial, archaeal, and fungal communities assemble in a differential manner and ask whether their assembly processes, community dynamics, and adaptability change over time as incipient soil develops into topsoil and subsoil. Furthermore, the results demonstrate that while community structure had spatial associations, and assembly processes were spatially structured, the dominant predictive functions were generally spatially homogeneous in the oligtrophic soil system. This trend has been observed in environmental and human microbiomes, including conserved functional gene abundances between desert and nondesert soil microbial communities in cross-biome metagenomic analyses (Fierer et al., 2012), functional convergence of the North American soil mycobiome (Talbot et al., 2014), and conserved functional gene profiles in the human gut microbiome (Lozupone et al., 2012),

4.1. Microbial Life Under Incipient Soil-Forming Conditions Resolved at Submeter Scale

The spatially intensive sampling scheme revealed presence of strong depth-dependent community composition profiles with absence of length-dependent profiles. This suggests that even though there is directional hydrologic transport lengthwise and thus the potential for lengthwise gradients in geochemical conditions and microbial composition, such potential for lengthwise gradients was not realized or was overwhelmed by vertically structured physical and chemical gradients that drove vertical gradients in microbial composition. This poses a question about measurement and modeling of biogeochemical function and microbial ecology of complex environments both in terms of dimension and scale of variation. In fact, upon comparing our results to a recent review of biophysical processes supporting microbial diversity (Tecon & Or, 2017), we found that our study is the first of its kind to evaluate microbial community structure at the submeter scale in relation to drivers like water content and total nitrogen. Submeter sampling strategies are becoming increasingly relevant; we posit the need to evaluate dimensionality alongside microscale factors impacting soil microbial diversity (Vos et al., 2013)

4.2. Depth-Dependent Physicochemical Variations

Our spatially intensive sampling indicated that surficial layers developed environments for microbial community assembly that was distinct from deeper layers. We determined that depth significantly impacted the measured and modeled environmental variables (BD, EC, TC = IC, TN, Water content values). Increased BD at the last depth could be a result of compaction or accumulation finer grained materials. However, a parallel analysis (Supporting Information Figure S12) revealed no significant change in particle-size distribution with depth, which leads us to conclude that compaction of the soil led to the differences observed in BD. Relative to deeper layers (40–80 cm), surface layers (0–30 cm) were subjected to greater exposure to air, higher salt concentrations, greater nitrogen input, faster vertical infiltration, and shorter contact-time with water. Additionally, the middle layers (40–60 cm) had the longest water contact-time due to water draining from the bottom layers of the toe slope, while the central layers stayed saturated.

Different spatial patterns of flow and water content distribution were observed within miniLEO as a consequence of moisture states. On one hand, prolonged rainfall events drove the overall system to a wet state, where a saturated zone was formed, creating a horizontal flow pattern that drove water laterally toward the seepage face and out of the mesocosm. On the other hand, between rainfall events, the redistribution of water was dominantly vertical, driven by gravitational effects and evaporative demand imposed at the

surface. These conditions lead to higher water contents close to toe slope, which had lower elevation and saw water accumulating in that region before exiting the system. Therefore, toeslope sections contributed significantly to the variation of water content during such periods. During wet periods, where a water table is formed, a saturated zone is produced along the length of the mesocosm, resulting in no variability along this dimension. These observations are supported by a modeled geochemical flux analysis (Dontsova et al., 2009) of solid phase evolution in the bulk parent material, which predicted that surficial layers of basalt in unsaturated zones will likely be subjected to greater primary mineral dissolution and secondary mineral precipitation than the lower layers that lie in the saturated zones. We note that two important biogeochemical conditions were not evaluated in the system: redox conditions and oxygen inputs. Therefore, we cannot infer how these variables impacted microhabitats in the system, though the strong influence of saturated conditions on oxygen penetration during rainfall events likely altered the availability of oxygen for both biotic and abiotic reactions.

4.3. Taxonomic Composition

Evaluation of incipient basaltic material in miniLEO showed a surprising increase in community diversity and phylogenetic composition relative to both the original parent material and the irrigation water. Greater than 50% of the 15 most abundant phyla in the weathered basaltic material were not detected in either the original parent material or the irrigation water. The development of structured microbial communities in the mesocosm that are distinct from the native composition present initially are likely the result of combined effects of (i) endogenous seed microbial colonies, which may have existed in the basaltic media; (ii) continuous input of exogenous microbes during irrigation events; and (iii) atmospheric deposition. Quantitative PCR (qPCR) analysis of 16S rRNA copy number of bacterial 16S rRNA gene (Supporting Information Figure S13) estimated an average three log-fold increase as compared to the parent material. Since the parent material itself was low in nutrients (Pangle et al., 2015), we conclude that microbial community development was aided by autochthonous nutrients from primary mineral weathering and dissolution, and allochthonous nutrients being deposited with water and air.

Depth-dependent variations in composition and relative abundance within the phylogenetic profiles suggested a selective pressure of depth. The transport and transfer of cells and nutrients in the mesocosm due to water infiltration and established flow paths may contribute to these patterns; null model analyses support this inference (see below). A recent study using the same basaltic parent material in the larger LEO hillslopes reported coherent depth-dependent spatial variations in solution chemistry and between distinct topographic regions of a landscape-scale lysimeter (Pohlmann et al., 2016). Since the current study evaluated a mesocosm with dimensions of 2 m × 1 m × 0.5 m, future directives of this work will evaluate whether similar depth-dependent associations persist, while length-dependent associations develop when examining microbial community structure, function, and assembly at pedon-scale and landscape-scale topographies. This may provide us with a better understanding of scale and dimension of microbial diversity and identify the scales at which depth and length dependencies start to form or collapse in soils.

Phylogenetic analysis of OTUs revealed *Proteobacteria* to be the dominant phylum in the system, an observation reported commonly for soils (Janssen, 2006) and seafloor basalts (Cockell et al., 2009; Orcutt et al., 2015; Singer et al., 2015). However, the surface samples more closely resembled oligotrophic community composition, with high relative abundance of *Actinobacteria*, as previously observed in other studies (Neilson et al., 2017, 2012). Their filamentous growth may be important in weathering processes in miniLEO, as observed in invasion and colonization of basaltic glass by actinobacterial groups, with a predicted mechanism of multidirectional filamentous growth across the surface with branches invading pore spaces (Cockell et al., 2009), and abundance on veneers of rock surface (Ortiz et al., 2013). Actinobacterial phyla have also been reported to scavenge atmospheric hydrogen in dry and hot oligotrophic environments (Greening et al., 2014; Lynch et al., 2014) and could potentially be carbon-fixation strategy employed in miniLEO surficial layer. Surprisingly, photoautotrophic Cyanobacteria, generally found in basalt environments (Olsson-Francis et al., 2012; Orcutt et al., 2011, 2015; Singer et al., 2015) was not an abundant phylum in miniLEO. The glass-enclosed structure of Biosphere 2 may either have filtered out necessary UV wavelengths required for cyanobacterial growth or sensitivity to incoming UV radiation may have driven these groups to move to areas with less light penetration at the cost of reduced cellular growth and productivity (Bebout & Garcia-pichel, 1995). However, phylogenetic potential for chemolithoautotrophy was found.

For example, *Nitrospirae* are primarily associated with chemoautrotrophy (Lücker et al., 2010), and numerous autotrophic phylotypes have been identified within the *Proteobacteria*, *Chloroflexi*, and *Planctomycetes* phyla. The large relative abundance of *Euryarcheota* (second most abundant phylum) in our samples is unique since terrestrial basalt environments do not report abundant presence of this group and neither did we observe any in the parent material and irrigation water. *Thermoplasmata* was the only euryarcheal class identified in the data set. These class of organisms are acidophiles and known to participate in methylotrophic methanogenesis (Poulsen et al., 2013). Although the miniLEO system overall does not appear to be conducive for either of these environments (low pH and anaerobic conditions), the gradual increase in Thermoplasmata (from 2.6% in the surficial sample to 11.9% in Depth_70 samples) is suggestive of pockets of microenvironments for such metabolisms to thrive and/or likelihood of organisms with novel metabolic strategies. Additionally, the absence of *Euryarcheota* in the parent material and irrigation water suggests that air is most likely the source, as highlighted in a recent study (Wehking et al., 2018), which shows *Euryarcheota* as a dominant airborne archaea in the fine particulate matter. Overall, our study revealed a highly diverse bacterial population and comparatively less diverse archaeal population in the incipient terrestrial basalt soil system subjected to enhanced weathering processes.

The vertical structure of microbial community composition was not mirrored by the predicted functional gene distributions within the mesocosm. This indicates the likelihood of redundancy within the predicted functions that allows the environmental conditions to select for different communities that may have the potential to perform similar functional roles. The presence of well-classified nitrogen fixation and nitrification taxa, along with the absence of predicted N-fixation genes and nitrification genes, was paradoxical. One interpretation is that the PICRUSt analyses are insufficient to infer functional potential within soil environments. Alternately, the mesocosm may have novel-nitrogen fixers and nitrifiers unique to the incipient basalt system that are too distant from the fully sequenced taxa that PICRUSt relies on. In fact, recent studies with approaches like genome assembled from metagenomes (MAGS) have shown that many organisms only contain partial pathways for biogeochemical reactions, including nitrogen fixation. Results from Anantharaman et al. (2016) show that in an extensive study of thousands of microbial genomes, only a few organisms had the potential to conduct multiple sequential metabolic reactions. This suggests that as an environment changes (for example, the depth dependent transitions in miniLEO), different groups of microorganisms might coordinately carry out major biogeochemical cycles. Additionally, our understanding and knowledge of the functional potential of well-studied microbial taxa is also being expanded. For example, Anantharaman et al. (2018) showed that the number of traditional sulfate/sulfite reducing microbes have doubled, with genomes of more than a dozen bacterial and archaeal phyla (outside the original group of reducers) being identified as capable of performing this metabolic function.

These functional prediction results should be treated with caution since they solely rely on bacterial and archaeal 16S rRNA gene sequences, do not take into account horizontal gene transfer across the genomes of the members of most microbial communities, are largely dependent on the availability of annotated reference genomes, and do not measure gene expression. PICRUSt analysis was also performed on closed-reference OTUs and therefore does not consider uncultured taxa with novel gene families. With this caveat, the PICRUSt data present the hypothesis that the observed transitions in beta diversity with depth do not reflect community trait-based variations in carbon and nitrogen cycling capacity. This approach also presents functional hypotheses that provide starting points for follow-on biogeochemical, microbiological, and molecular studies aimed at understanding connections among physical, chemical, and biological features of the system.

4.4. Community Assembly

Community assembly processes operate over space and time and influence community structure and eventually microbial metabolisms (Graham, Knelman, et al., 2016). So far, attempts have been made to understand these processes in a broad range of biomes such as recently deglaciated soils (Castle et al., 2016; Nemergut et al., 2013), lake biofilms (Jackson et al., 2001), water pipes (Martiny et al., 2003), river sediments (Graham, Crump, et al., 2016; Stegen, Fredrickson, et al., 2016; Stegen, Konopka, et al., 2016; Stegen et al., 2015, 2012), subseafloor sediments (Starnawski et al., 2017), and global desert soils (Caruso et al., 2011). Our study examined a unique system where community assembly processes in an incipient terrestrial system were assessed in a highly controlled and spatially explicit manner.

While the influence of variable selection did increase toward the upper layers, there was significant spatial heterogeneity in its influence, consistent with previous work examining the spatial structure of assembly processes (Stegen et al., 2015). This is particularly interesting given that miniLEO started with a relatively homogeneous environment, which may be expected to lead to strong influences of homogeneous selection. The lack of βNTI values <-2 across miniLEO, however, indicates that homogeneous selection was not influential in miniLEO.

Variable selection clearly dominated the system, but spatial processes also appeared to have some influence. In particular, the influence of homogenizing dispersal was found to have nonrandom spatial structure, with the strongest influences found near the center of miniLEO. This spatial structure makes intuitive sense in that dispersing organisms can enter locations within the central part of miniLEO from any direction, whereas along the edges dispersal is not possible along one or more sides of a given voxel. In addition, the central part of miniLEO maintained saturated conditions for a longer period of time than the surficial or deepest layers. Saturated conditions likely increased the ability of microorganisms to cross large pore spaces, thereby facilitating dispersal (Ebrahimi & Or, 2014). We infer that access to a voxel from all sides in addition to well-connected pores facilitated dispersal into voxels found throughout the central region of miniLEO, thereby increasing the relative contribution of homogenizing dispersal. These inferences are conceptually consistent with previous work finding greater influences of homogenizing dispersal in a more hydrologically permeable geological formation (Stegen et al., 2013). While we do not have time series data for community composition, we infer that the abiotic environment evolved through time in a spatially heterogeneous manner that resulted in localized selection pressures, thereby leading to strong influences of variable selection. While logistically difficult, it would be fascinating to study temporal dynamics in the spatial structure of variable selection, and other community assembly processes. To the best of our knowledge, such spatiotemporal characterization has not been attempted in any system.

5. Conclusion

This study is a first attempt to characterize an incipient basalt soil weathering system using a combination of phylogenetic information and a null model approach. Community diversity metrics, phylogenetic profiles, and the null model approach indicate that assembly was governed by deterministic processes and was a function of depth-dependent microenvironments developing as a result of localized abiotic drivers. Preliminary predictive functional signatures did not differentiate with depth suggesting that redundancy in potential function may be present among the diverse phyla identified throughout the mesocosm. Nonetheless, water content explained a significant amount of variation in phylogenetic composition, suggesting that the distribution of soil moisture—including the relative prevalence of saturated versus unsaturated conditions—was governing the structure, composition, and assembly of the communities. The null model analyses in particular show that community composition was governed by deterministic processes (e.g., selection by the abiotic environment) and that transport processes (i.e., dispersal) had relatively little influence.

The broader implications of our results lie in assessing the length scales of variation in physical, chemical, and biological features between vertical and horizontal dimensions in natural environments. The results here suggest that for oligotrophic systems there are major differences in the length scales of variation between vertical and horizontal dimensions for which limited biogeochemical activity results in weak horizontal gradients along hydrologic flow paths. We hypothesize that in more productive/active systems, the length scales of variation between vertical and horizontal dimensions may converge due to the potential for more biogeochemical activity along horizontal flow paths causing rapid shifts in environmental conditions.

Because our results and inferences capture one snapshot after 2 years within a dynamic system, they are likely to change and present a different perspective on incipient soil development as the system continues to mature. Extending our spatially explicit characterization to be fully spatiotemporal and capture vertical and horizontal dimensions of hydrobiogeochemical signatures represents an exciting new frontier that has the potential to reveal key feedbacks between community assembly processes, microbial function, and (bio)geochemical processes. This can potentially be used to develop predictive understanding of marginal

soil progression (e.g., transformation of marginal soils, soils recovering from disturbances such as fire, erosion, intensive agricultural practices, or contamination, and soil material exposed by retreating glaciers).

Acknowledgments
AS, JN, AAMN, YW, EH, TT, KD, JC, PT, and RM wish to acknowledge support of NSF-funded projects EAR-1344552, EAR-1340912, EAR-1417097, and Philecology Foundation of Fort Worth Texas. AAMN would like to acknowledge the support received by the Brazilian Scientific Mobility Program promoted by CAPES. Additional funding support were provided by the Water, Environmental, and Energy Solutions (WEES) initiative at the University of Arizona and by the office of Research, Discovery and Innovation's Accelerate for Success Grant at the University of Arizona. We would also like to acknowledge Daniel Laubitz at the University of Arizona Genomics Core for method development and sequencing of low-template samples. JCS was supported by the US Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of Subsurface Biogeochemical Research Program's Scientific Focus Area (SFA) at Pacific Northwest National Laboratory (PNNL). PNNL is operated for DOE by Battelle Memorial Institute under contract DE-AC06-76RLO 1830. The authors declare that they have no conflicts of interest. Conforming with the AGU data policy, data are available from the supporting information. Supporting data as two tables, three supporting datasets, and eleven figures are included in the supporting information file, sequence data is deposited at Sequence Read Archive SRP116044 (Accession PRJNA392820), and R code for Null Modeling can be found at https://github.com/stegen/Stegen_etal_ISME_2013.

References

Ackerly, D. D. (2004). Adaptation, niche conservatism, and convergence: Comparative studies of leaf evolution in the California chaparral. *The American Naturalist*, *163*(5), 654–671. https://doi.org/10.1086/383062

Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., et al. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nature Communications*, *7*(1). https://doi.org/10.1038/ncomms13219

Anantharaman, K., Hausmann, B., Jungbluth, S. P., Kantor, R. S., Lavy, A., Warren, L. A., et al. (2018). Expanded diversity of microbial groups that shape the dissimilatory sulfur cycle. *The ISME Journal*, *12*(7), 1715–1728. https://doi.org/10.1038/s41396-018-0078-0

Banfield, J. F., Barker, W. W., Welch, S. A., & Taunton, A. (1999). Biological impact on mineral dissolution: Application of the lichen model to understanding mineral weathering in the rhizosphere. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(7), 3404–3411. https://doi.org/10.1073/PNAS.96.7.3404

Bazzaz, F. A. (1991). Habitat selection in plants. *The American Naturalist*, *137*, S116–S130. https://doi.org/10.1086/285142

Bebout, B. M., & Garcia-pichel, F. (1995). UV B-induced vertical migrations of cyanobacteria in a microbial mat. *Applied and Environmental Microbiology*, *61*(12), 4215–4222. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1388643/pdf/hw4215.pdf

Bell, G. (2001). Neutral macroecology. *Science*, *293*(5539), 2413–2418. https://doi.org/10.1126/science.293.5539.2413

Bengtson, S., Ivarsson, M., Astolfo, A., Belivanova, V., Broman, C., Marone, F., & Stampanoni, M. (2014). Deep-biosphere consortium of fungi and prokaryotes in Eocene subseafloor basalts. *Geobiology*, *12*(6), 489–496. https://doi.org/10.1111/gbi.12100

Bockheim, J. G., Gennadiyev, A. N., Hartemink, A. E., & Brevik, E. C. (2014). Soil-forming factors and soil taxonomy. *Geoderma*, *226-227*, 231–237. https://doi.org/10.1016/j.geoderma.2014.02.016

Bradley, J. A., Singarayer, J. S., & Anesio, A. M. (2014). Microbial community dynamics in the forefield of glaciers. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1795). Retrieved from http://rspb.royalsocietypublishing.org/content/281/1795/20140882, https://doi.org/10.1098/rspb.2014.0882

Brown, S. P., & Jumpponen, A. (2014). Contrasting primary successional trajectories of fungi and bacteria in retreating glacier soils. *Molecular Ecology*, *23*(2), 481–497. https://doi.org/10.1111/mec.12487

Bu, C., Wang, Y., Ge, C., Ahmad, H. A., Gao, B., & Ni, S.-Q. (2017). Dissimilatory nitrate reduction to ammonium in the Yellow River Estuary: Rates, abundance, and Community Diversity. *Scientific Reports*, *7*(1), 6830. https://doi.org/10.1038/s41598-017-06404-8

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., & Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(34), 14,288–14,293. https://doi.org/10.1073/pnas.1101591108

Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., & Knight, R. (2010). PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics*, *26*(2), 266–267. https://doi.org/10.1093/bioinformatics/btp636

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335–336. https://doi.org/10.1038/nmeth.f.303

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, *6*(8), 1621–1624. https://doi.org/10.1038/ismej.2012.8

Caruso, T., Chan, Y., Lacap, D. C., Lau, M. C. Y., McKay, C. P., & Pointing, S. B. (2011). Stochastic and deterministic processes interact in the assembly of desert microbial communities on a global scale. *The ISME Journal*, *5*(9), 1406–1413. https://doi.org/10.1038/ismej.2011.21

Castle, S. C., Nemergut, D. R., Grandy, A. S., Leff, J. W., Graham, E. B., Hood, E., et al. (2016). Biogeochemical drivers of microbial community convergence across actively retreating glaciers. *Soil Biology and Biochemistry*, *101*, 74–84. https://doi.org/10.1016/j.soilbio.2016.07.010

Cavender-Bares, J., Ackerly, D. D., Baum, D. A., & Bazzaz, F. A. (2004). Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist*, *163*(6), 823–843. https://doi.org/10.1086/386375

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, *12*(7), 693–715. https://doi.org/10.1111/j.1461-0248.2009.01314.x

Cockell, C. S., Olsson, K., Knowles, F., Kelly, L., Herrera, A., Thorsteinsson, T., & Marteinsson, V. (2009). Bacteria in weathered basaltic glass, Iceland. *Geomicrobiology Journal*, *26*(7), 491–507. https://doi.org/10.1080/01490450903061101

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, *72*(7), 5069–5072. https://doi.org/10.1128/AEM.03006-05

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., & Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(11), E1326–E1332. https://doi.org/10.1073/pnas.1414261112

Donoghue, M. J. (2008). Colloquium paper: A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(Supplement 1), 11,549–11,555. https://doi.org/10.1073/pnas.0801962105

Dontsova, K., Steefel, C. I., Desilets, S., Thompson, A., & Chorover, J. (2009). Solid phase evolution in the Biosphere 2 hillslope experiment as predicted by modeling of hydrologic and geochemical fluxes. *Hydrology and Earth System Sciences*, *13*(12), 2273–2286. https://doi.org/10.5194/hess-13-2273-2009

Dumbrell, A. J., Nelson, M., Helgason, T., Dytham, C., & Fitter, A. H. (2010). Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME Journal*, *4*(3), 337–345. https://doi.org/10.1038/ismej.2009.122

Ebrahimi, A. N., & Or, D. (2014). Microbial dispersal in unsaturated porous media: Characteristics of motile bacterial cell motions in unsaturated angular pore networks. *Water Resources Research*, *50*, 7406–7429. https://doi.org/10.1002/2014WR015897

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Emerson, B. C., & Gillespie, R. G. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends in Ecology & Evolution*, *23*(11), 619–630. https://doi.org/10.1016/j.tree.2008.07.005

Enquist, B. J., Sanderson, J., & Weiser, M. D. (2002). Modeling macroscopic patterns in ecology. *Science*, *295*(5561), 1835–1837. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11890180

Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S., & Thomas, T. (2012). Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(27), E1878–E1887. https://doi.org/10.1073/pnas.1203287109

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., et al. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21,390–21,395. https://doi.org/10.1073/pnas.1215210110

Fine, P. V. A., Miller, Z. J., Mesones, I., Irazuzta, S., Appel, H. M., Stevens, M. H. H., et al. (2006). The growth–defense trade-off and habitat specialization by plants in Amazonian forests. *Ecology*, *87*(sp7), S150–S162. https://doi.org/10.1890/0012-9658(2006)87[150:TGTAHS]2.0.CO;2

Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature*, *459*(7244), 193–199. https://doi.org/10.1038/nature08058

Gillespie, R. (2004). Community assembly through adaptive radiation in Hawaiian spiders. *Science*, *303*(5656), 356–359. Retrieved from http://science.sciencemag.org/content/303/5656/356, https://doi.org/10.1126/science.1091875

Gittel, A., BÃirta, J., KohoutovÃi, I., Schnecker, J., Wild, B., ÄŒapek, P., et al. (2014). Site- and horizon-specific patterns of microbial community structure and enzyme activities in permafrost-affected soils of Greenland. *Frontiers in Microbiology*, *5*, 541. https://doi.org/10.3389/fmicb.2014.00541

Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: Linking ecological and evolutionary processes across space in time. *Ecology Letters*, *11*(12), 1265–1277. https://doi.org/10.1111/j.1461-0248.2008.01256.x

Graham, E. B., Crump, A. R., Resch, C. T., Fansler, S., Arntzen, E., Kennedy, D. W., et al. (2016). Coupling spatiotemporal community assembly processes to changes in microbial metabolism. *Frontiers in Microbiology*, *7*(1949). https://doi.org/10.3389/fmicb.2016.01949

Graham, E. B., Knelman, J. E., Schindlbacher, A., Siciliano, S., Breulmann, M., Yannarell, A., et al. (2016). Microbes as engines of ecosystem function: When does community structure enhance predictions of ecosystem processes? *Frontiers in Microbiology*, *7*(214). https://doi.org/10.3389/fmicb.2016.00214

Green, J. L., Holmes, A. J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., et al. (2004). Spatial scaling of microbial eukaryote diversity. *Nature*, *432*(7018), 747–750. https://doi.org/10.1038/nature03034

Greening, C., Berney, M., Hards, K., Cook, G. M., & Conrad, R. (2014). A soil actinobacterium scavenges atmospheric H2 using two membrane-associated, oxygen-dependent [NiFe] hydrogenases. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(11), 4257–4261. https://doi.org/10.1073/pnas.1320586111

Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., et al. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research*, *21*(3), 494–504. https://doi.org/10.1101/gr.112730.110

Hansel, C. M., Fendorf, S., Jardine, P. M., & Francis, C. A. (2008). Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile. *Applied and Environmental Microbiology*, *74*(5), 1620–1633. https://doi.org/10.1128/AEM.01787-07

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*(7), 497–506. https://doi.org/10.1038/nrmicro2795

Hartmann, M., Lee, S., Hallam, S. J., & Mohn, W. W. (2009). Bacterial, archaeal and eukaryal community structures throughout soil horizons of harvested and naturally disturbed forest stands. *Environmental Microbiology*, *11*(12), 3045–3062. https://doi.org/10.1111/j.1462-2920.2009.02008.x

He, F., & Gaston, K. J. (2003). Occupancy, spatial variance, and the abundance of species. *The American Naturalist*, *162*(3), 366–375. https://doi.org/10.1086/377190

Horner-Devine, M. C., Lage, M., Hughes, J. B., & Bohannan, B. J. M. (2004). A taxa–area relationship for bacteria. *Nature*, *432*(7018), 750–753. https://doi.org/10.1038/nature03073

Hubbell, S. P. (2006). Neutraltheoryandtheevolutionofecologicalequivalence. *Ecology*, *87*(6), 1387–1398. https://doi.org/10.1890/0012-9658(2006)87[1387:NTATEO]2.0.CO;2

Jackson, C. R., Churchill, P. F., & Roden, E. E. (2001). Successional changes in bacterial assemblage structure during epilithic biofilm development. *Ecology*, *82*(2), 555. https://doi.org/10.2307/2679879

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and Environmental Microbiology*, *72*(3), 1719–1728. https://doi.org/10.1128/AEM.72.3.1719-1728.2006

Jenny, H. K. (1941). Factors of soil formation. *Soil Science*. *52*(5), 415. Retrieved from https://books.google.com/books?hl=en&lr=&id=orjZZS3H-hAC&oi=fnd&pg=PP1&dq=soil+formation&ots=fInIh2aRmg&sig=51Beiw2aQIL81T6A3lSUr9Lc1Ng#v=onepage&q=soil, https://doi.org/10.1097/00010694-194111000-00009

Jonathan Davies, T., Meiri, S., Barraclough, T. G., & Gittleman, J. L. (2007). Species co-existence and character divergence across carnivores. *Ecology Letters*, *10*(2), 146–152. https://doi.org/10.1111/j.1461-0248.2006.01005.x

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, *45*(D1), D353–D361. https://doi.org/10.1093/nar/gkw1092

Kim, M., Pangle, L. A., Cardoso, C., Lora, M., Volkmann, T. H. M., Wang, Y., et al. (2016). Transit time distributions and StorAge selection functions in a sloping soil lysimeter with time-varying flow paths: Direct observation of internal and external transport variability. *Water Resources Research*, *52*, 7105–7129. https://doi.org/10.1002/2016WR018620

Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814–821. https://doi.org/10.1038/nbt.2676

Laubitz, D., Harrison, C. A., Midura-Kiela, M. T., Ramalingam, R., Larmonier, C. B., Chase, J. H., et al. (2016). Reduced epithelial Na^+/H^+ exchange drives gut microbial dysbiosis and promotes inflammatory response in T cell-mediated murine colitis. *PLoS ONE*, *11*(4), e0152044. https://doi.org/10.1371/journal.pone.0152044

Loreau, M. (2004). Does functional redundancy exist? *Oikos*, *104*(3), 606–611. https://doi.org/10.1111/j.0030-1299.2004.12685.x

Lozupone, C., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, *489*(7415), 220–230. https://doi.org/10.1038/nature11550

Lücker, S., Wagner, M., Maixner, F., Pelletier, E., Koch, H., Vacherie, B., et al. (2010). A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proceedings of the National Academy of Sciences*, *107*(30), 13479–13484. https://doi.org/10.1073/pnas.1003860107

Lynch, R. C., Darcy, J. L., Kane, N. C., Nemergut, D. R., & Schmidt, S. K. (2014). Metagenomic evidence for metabolism of trace atmospheric gases by high-elevation desert Actinobacteria. *Frontiers in Microbiology*, *5*(698). https://doi.org/10.3389/fmicb.2014.00698

Mapelli, F., Marasco, R., Balloi, A., Rolli, E., Cappitelli, F., Daffonchio, D., & Borin, S. (2012). Mineral–microbe interactions: Biotechnological potential of bioweathering. *Journal of Biotechnology*, *157*(4), 473–481. https://doi.org/10.1016/j.jbiotec.2011.11.013

Martiny, A. C., Jørgensen, T. M., Albrechtsen, H.-J., Arvin, E., & Molin, S. (2003). Long-term succession of structure and diversity of a biofilm formed in a model drinking water distribution system. *Applied and Environmental Microbiology*, *69*(11), 6899–6907. https://doi.org/10.1128/AEM.69.11.6899-6907.2003

McMurdie, P. J., Holmes, S., Kindt, R., Legendre, P., & O'Hara, R. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Michel, H. M., & Williams, M. A. (2011). Soil habitat and horizon properties impact bacterial diversity and composition. *Soil Science Society of America Journal*, *75*(4), 1440–1448. https://doi.org/10.2136/sssaj2010.0171

Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., et al. (2007). Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecology Letters*, *10*(4), 315–331. https://doi.org/10.1111/j.1461-0248.2007.01020.x

Neilson, J. W., Califf, K., Cardona, C., Copeland, A., van Treuren, W., Josephson, K. L., et al. (2017). Significant impacts of increasing aridity on the arid soil microbiome. *MSystems*, *2*(3). Retrieved from http://msystems.asm.org/content/2/3/e00195-16, https://doi.org/10.1128/mSystems.00195-16

Neilson, J. W., Quade, J., Ortiz, M., Nelson, W. M., Legatzki, A., Tian, F., et al. (2012). Life at the hyperarid margin: Novel bacterial diversity in arid soils of the Atacama Desert, Chile. *Extremophiles*, *16*(3), 553–566. https://doi.org/10.1007/s00792-012-0454-z

Nelson, M. B., Martiny, A. C., & Martiny, J. B. H. (2016). Global biogeography of microbial nitrogen-cycling traits in soil. *Proceedings of the National Academy of Sciences*, *113*(29), 8033–8040. https://doi.org/10.1073/pnas.1601070113

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews: MMBR*, *77*(3), 342–356. https://doi.org/10.1128/MMBR.00051-12

Oksanen, J. (2015). Multivariate analysis of ecological communities in R: Vegan tutorial. Retrieved from http://cc.oulu.fi/~jarioksa/opetus/metodi/vegantutor.pdf

Olsson-Francis, K., Simpson, A. E., Wolff-Boenisch, D., & Cockell, C. S. (2012). The effect of rock composition on cyanobacterial weathering of crystalline basalt and rhyolite. *Geobiology*, *10*(5), 434–444. https://doi.org/10.1111/j.1472-4669.2012.00333.x

Orcutt, B. N., Sylvan, J. B., Knab, N. J., & Edwards, K. J. (2011). Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiology and Molecular Biology Reviews: MMBR*, *75*(2), 361–422. https://doi.org/10.1128/MMBR.00039-10

Orcutt, B. N., Sylvan, J. B., Rogers, D. R., Delaney, J., Lee, R. W., & Girguis, P. R. (2015). Carbon fixation by basalt-hosted microbial communities. *Frontiers in Microbiology*, *6*(904). https://doi.org/10.3389/fmicb.2015.00904

Ortiz, M., Legatzki, A., Neilson, J. W., Fryslie, B., Nelson, W. M., Wing, R. a., et al. (2013). Making a living while starving in the dark: Metagenomic insights into the energy dynamics of a carbonate cave. *The ISME Journal*, *8*(2), 478–491. https://doi.org/10.1038/ismej.2013.159

Pangle, L. A., DeLong, S. B., Abramson, N., Adams, J., Barron-Gafford, G. A., Breshears, D. D., et al. (2015). The Landscape Evolution Observatory: A large-scale controllable infrastructure to study coupled Earth-surface processes. *Geomorphology*, *244*, 190–203. https://doi.org/10.1016/j.geomorph.2015.01.020

Pangle, L. A., Kim, M., Cardoso, C., Lora, M., Meira Neto, A. A., Volkmann, T. H. M., et al. (2017). The mechanistic basis for storage-dependent age distributions of water discharged from an experimental hillslope. *Water Resources Research*, *53*, 2733–2754. https://doi.org/10.1002/2016WR019901

Pennington, R. T., Richardson, J. E., & Lavin, M. (2006). Insights into the historical construction of species-rich biomes from dated plant phylogenies, neutral ecological theory and phylogenetic community structure. *New Phytologist*, *172*(4), 605–616. https://doi.org/10.1111/j.1469-8137.2006.01902.x

Pholchan, M. K., Baptista, J. d. C., Davenport, R. J., Sloan, W. T., & Curtis, T. P. (2013). Microbial community assembly, theory and rare functions. *Frontiers in Microbiology*, *4*(68). https://doi.org/10.3389/fmicb.2013.00068

Pohlmann, M., Dontsova, K., Root, R., Ruiz, J., Troch, P., & Chorover, J. (2016). Pore water chemistry reveals gradients in mineral transformation across a model basaltic hillslope. *Geochemistry, Geophysics, Geosystems*, *17*, 2054–2069. https://doi.org/10.1002/2016GC006270

Poulsen, M., Schwab, C., Borg Jensen, B., Engberg, R. M., Spang, A., Canibe, N., et al. (2013). Methylotrophic methanogenic Thermoplasmata implicated in reduced methane emissions from bovine rumen. *Nature Communications*, *4*(1), 1428. https://doi.org/10.1038/ncomms2432

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, *5*(3), e9490. https://doi.org/10.1371/journal.pone.0009490

R Core Team (2014). R: A language and environment for statistical computing. 2014, R Foundation for Statistical Computing, Vienna, Austria. Retrieved April 17, 2017, from http://www.r-project.org/

Ricklefs, R. E. (2004). A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, *7*(1), 1–15. https://doi.org/10.1046/j.1461-0248.2003.00554.x

Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., & Etienne, R. S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, *27*(4), 203–208. https://doi.org/10.1016/j.tree.2012.01.004

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schmidt, S. K., Nemergut, D. R., Darcy, J. L., & Lynch, R. (2014). Do bacterial and fungal communities assemble differently during primary succession? *Molecular Ecology*, *23*(2), 254–258. https://doi.org/10.1111/mec.12589

Sengupta, A., Wang, Y., Meira Neto, A. A., Matos, K. A., Dontsova, K., Root, R., et al. (2016). Soil lysimeter excavation for coupled hydrological, geochemical, and microbiological investigations. *Journal of Visualized Experiments*, *2016*(115). https://doi.org/10.3791/54536

Seuradge, B. J., Oelbermann, M., & Neufeld, J. D. (2017). Depth-dependent influence of different land-use systems on bacterial biogeography. *FEMS Microbiology Ecology*, *93*(2), fiw239. https://doi.org/10.1093/femsec/fiw239

Singer, E., Chong, L. S., Heidelberg, J. F., & Edwards, K. J. (2015). Similar microbial communities found on two distant seafloor basalts. *Frontiers in Microbiology*, *6*(1409). https://doi.org/10.3389/fmicb.2015.01409

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., & Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology*, *8*(4), 732–740. https://doi.org/10.1111/j.1462-2920.2005.00956.x

Sloan, W. T., Woodcock, S., Lunn, M., Head, I. M., & Curtis, T. P. (2007). Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microbial Ecology*, *53*(3), 443–455. https://doi.org/10.1007/s00248-006-9141-x

Slobodkina, G. B., Mardanov, A. V., Ravin, N. V., Frolova, A. A., Chernyh, N. A., Bonch-Osmolovskaya, E. A., & Slobodkin, A. I. (2017). Respiratory ammonification of nitrate coupled to anaerobic oxidation of elemental sulfur in deep-sea autotrophic thermophilic bacteria. *Frontiers in Microbiology*, *8*(87). https://doi.org/10.3389/fmicb.2017.00087

Starnawski, P., Bataillon, T., Ettema, T. J. G., Jochum, L. M., Schreiber, L., Chen, X., et al. (2017). Microbial community assembly and evolution in subseafloor sediment. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(11), 2940–2945. https://doi.org/10.1073/pnas.1614190114

Stegen, J. C., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E., Nelson, W. C., Arntzen, E. V., et al. (2016). Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. *Nature Communications*, *7*(1), 11237. https://doi.org/10.1038/ncomms11237

Stegen, J. C., Konopka, A., McKinley, J. P., Murray, C., Lin, X., Miller, M. D., et al. (2016). Coupling among microbial communities, biogeochemistry, and mineralogy across biogeochemical facies. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep30553

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, *7*(11), 2069–2079. https://doi.org/10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Fredrickson, J. K., & Konopka, A. E. (2015). Estimating and mapping ecological processes influencing microbial community assembly. *Frontiers in Microbiology*, *6*(370). https://doi.org/10.3389/fmicb.2015.00370

Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME Journal*, *6*(9), 1653–1664. https://doi.org/10.1038/ismej.2012.22

Strauss, S. Y., Webb, C. O., & Salamin, N. (2006). Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(15), 5841–5845. https://doi.org/10.1073/pnas.0508073103

Swenson, N. G., Enquist, B. J., Thompson, J., & Zimmerman, J. K. (2007). The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities. *Ecology*, *88*(7), 1770–1780. https://doi.org/10.1890/06-1499.1

Talbot, J. M., Bruns, T. D., Taylor, J. W., Smith, D. P., Branco, S., Glassman, S. I., et al. (2014). Endemism and functional convergence across the North American soil mycobiome. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(17), 6341–6346. https://doi.org/10.1073/pnas.1402584111

Tecon, R., & Or, D. (2017). Biophysical processes supporting the diversity of microbial life in soil. *FEMS Microbiology Reviews*, *41*(5), 599–623. https://doi.org/10.1093/femsre/fux039

Thorup, C., Schramm, A., Findlay, A. J., Finster, K. W., & Schreiber, L. (2017). Disguised as a sulfate reducer: Growth of the Deltaproteobacterium Desulfurivibrio alkaliphilus by sulfide oxidation with nitrate. *MBio*, *8*(4), e00671–e00617. https://doi.org/10.1128/mBio.00671-17

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(30), 10,854–10,861. https://doi.org/10.1073/pnas.0403458101

Turner, S., Mikutta, R., Meyer-Stüve, S., Guggenberger, G., Schaarschmidt, F., Lazar, C. S., et al. (2017). Microbial community dynamics in soil depth profiles over 120,000 years of ecosystem development. *Frontiers in Microbiology*, *8*(874). https://doi.org/10.3389/fmicb.2017.00874

Vamosi, S. M., Heard, S. B., Vamosi, J. C., & Webb, C. O. (2009). Emerging patterns in the comparative analysis of phylogenetic community structure. *Molecular Ecology*, *18*(4), 572–592. https://doi.org/10.1111/j.1365-294X.2008.04001.x

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS Microbiology Reviews*, *37*(6), 936–954. https://doi.org/10.1111/1574-6976.12023

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*(16), 5261–5267. https://doi.org/10.1128/AEM.00062-07

Wani, A. A., Surakasi, V. P., Siddharth, J., Raghavan, R. G., Patole, M. S., Ranade, D., & Shouche, Y. S. (2006). Molecular analyses of microbial diversity associated with the Lonar soda lake in India: An impact crater in a basalt area. *Research in Microbiology*, *157*(10), 928–937. https://doi.org/10.1016/j.resmic.2006.08.005

Warenes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., et al. (2016). Various R Programming Tools for Plotting Data. Retrieved from https://cran.r-project.org/web/packages/gplots/gplots.pdf

Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, *33*(1), 475–505. https://doi.org/10.1146/annurev.ecolsys.33.010802.150448

Wehking, J., Pickersgill, D. A., Bowers, R. M., Teschner, D., Pöschl, U., Fröhlich-Nowoisky, J., & Després, V. R. (2018). Community composition and seasonal changes of archaea in coarse and fine air particulate matter. *Biogeosciences*, *15*(13), 4205–4214. https://doi.org/10.5194/bg-15-4205-2018

Weiher, E. (1999). *Ecological assembly rules perspectives, advances, retreats*. Cambridge University Press. https://doi.org/10.1017/CBO9780511542237

Weil, R. R., & Brady, N. C. (2017). *The nature and properties of soils*, (15th ed.). New Jersey: Pearson.

Will, C., Thurmer, A., Wollherr, A., Nacke, H., Herold, N., Schrumpf, M., et al. (2010). Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. *Applied and Environmental Microbiology*, *76*(20), 6751–6759. https://doi.org/10.1128/AEM.01063-10

Woodcock, S., Van Der Gast, C. J., Bell, T., Lunn, M., Curtis, T. P., Head, I. M., & Sloan, W. T. (2007). Neutral assembly of bacterial communities. *FEMS Microbiology Ecology*, *62*(2), 171–180. https://doi.org/10.1111/j.1574-6941.2007.00379.x

Erratum

In the originally published version of this article, the following authors were erroneously omitted from the author byline: Triffon Tatarin, Edward Hunt, and Katerina Dontsova. As a result, the Author Contributions and Acknowledgments sections published incorrectly. The author byline, contributions, and acknowledgments have since been corrected, and this version may be considered the authoritative version of record.